



Data Science Checklist for AIOps Applications

Augmented Intelligence, machine learning, and analytics are increasingly deployed in service management systems and tool sets to enhance their performance. How and when they are used separately and in combination defines and/or limits how effective the AIOps application will be in improving service assurance processes from fault to customer experience management.

INDUSTRY NARRATIVE

Next generation, wide scope AIOps applications should employ:

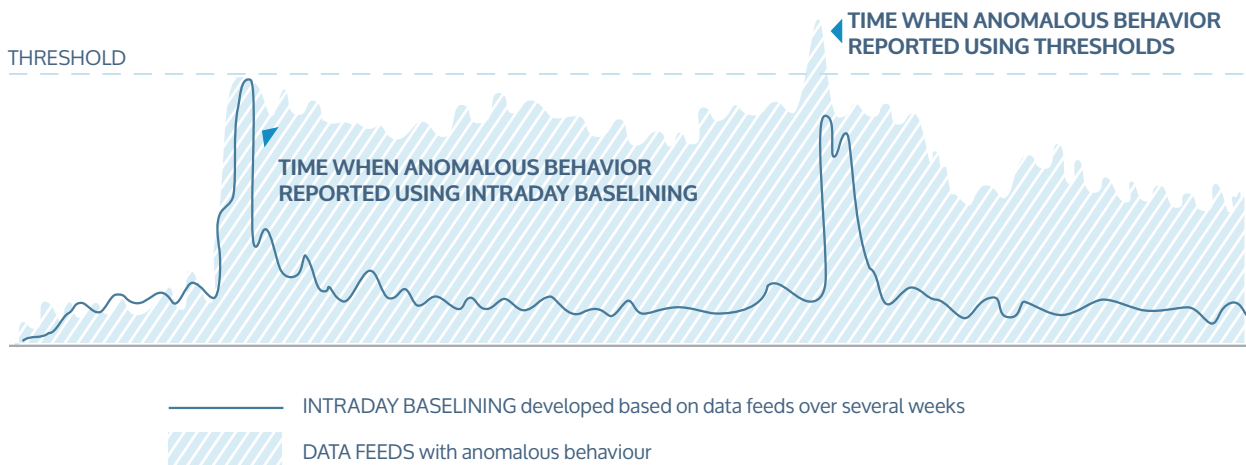
- 1 **Advanced anomaly detection** to enable adoption to changes in seasonality, magnitude, and deviation over short intraday time periods
- 2 **Stochastic models** to reduce noise and detect performance issues and faults early
- 3 **Affinity analysis** beyond simple temporal correlation to identify related events and define the root cause
- 4 **Probable cause** determination using severity, entropy, and eccentricity metrics for every dimension to distinguish between symptoms and root cause
- 5 **Ontology reasoning** to optimize performance management

This paper will look at each of these data science capabilities, describe them and define how and why they are important in service performance management.

ADVANCED ANOMALY DETECTION

Simple threshold-based anomaly detection simply does not work well in modern data centers or complex networks due to rapidly changing workloads and volumes, regardless of whether the thresholds are user-set or statistically learned. They are likely to trigger false positives during peak usage and heavy loads and miss true positives during quieter periods. Instead, more adaptive anomaly detection is required, one that continuously learns seasonality in load and usage, and triggers alerts based on deviation from expected behavior.

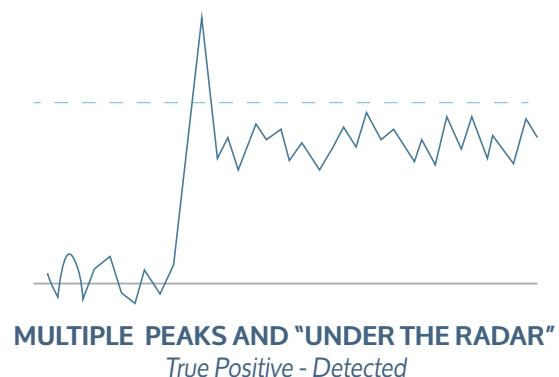
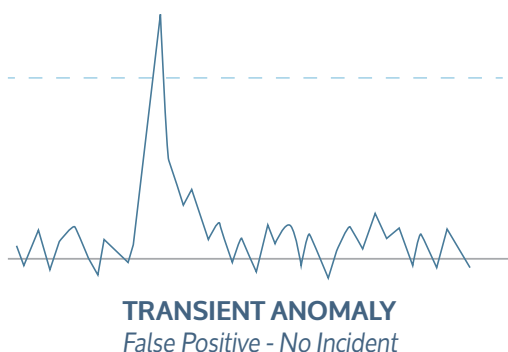
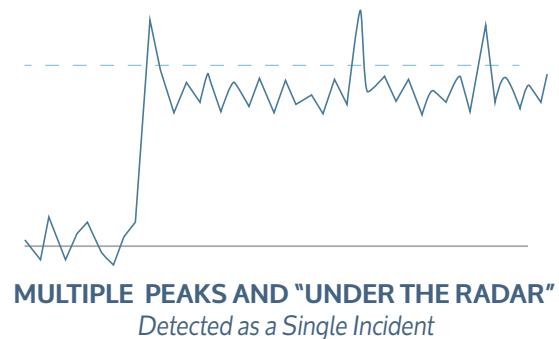
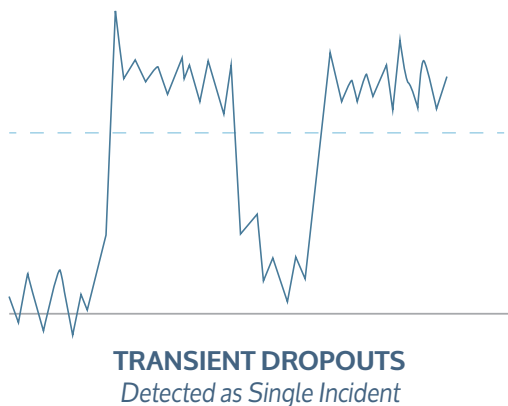
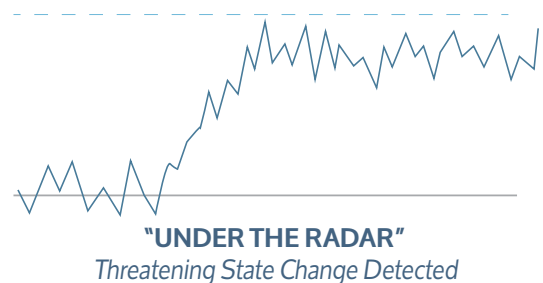
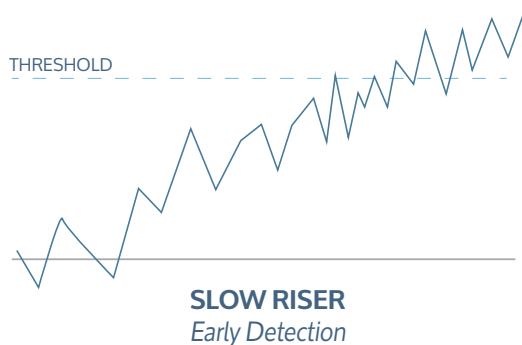
Utilizing unsupervised machine learning, advanced anomaly detection relies on learning time-varying baselines on each metric and dimension as data is ingested, and continuously updating them as more data is collected. Triggering alerts based on deviations from learned baselines provides more robust alerting, i.e., capturing the significant anomalies occurring during low usages time periods, while reducing the false positive noise that often occurs peak periods.



STOCHASTIC MODELS

Anomalous signals arising from the various event and metric streams being monitored are often transient, resulting from temporary usage spikes or statistical noise. These transient anomalies do not necessarily indicate a persistent problem. The ability to identify anomalies that are both significant and non-transient enables operations teams to focus on those problems that truly need fixes, and hence improves operational efficiency.

Stochastic models excel at separating signal from noise. For this reason, they are widely used on Wall Street to model the seemingly random fluctuations in market behavior and volatility, and predict when market conditions have changed. And for similar reasons, they are useful in the noisy world of data centers and IT operations. These models can continuously monitor and evaluate the behavior of every metric, event and entity looking for non-transient anomalies and suspicious changes in state that indicate an "incident" is occurring and needs correcting. Stochastic models correctly detect the patterns that other techniques will typically misclassify, identify late, or miss altogether.



Stochastic models also allow for a dynamic “look-back” period to capture the point in time where the system first exhibited a detrimental change in behavior. This look-back period is likely to spot issues before the signal is declared to be an “incident” by most fault and performance management systems. Look back periods can even identify “slow risers” where the change in behavior takes a long time to manifest into a service-impacting Incident.

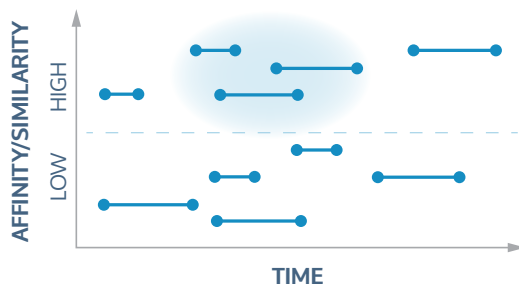


Dynamic look-back windows based on behavioral changes are more likely to be able to properly identify and correlate early triggering service events.

AFFINITY – BEYOND TEMPORAL CORRELATION

Most fault and performance management systems group anomalous signals occurring during the same time period. But most do not analyze and group based on the similarity of system components and dimension values (features). VIA AIOps goes further by also grouping based on an affinity score between anomalous signals, based on statistically rigorous similarity measures, such as the Jaccard Similarity Index. Grouping in this manner combined with temporal correlation can better determine if anomalies are related and if they should be grouped together or treated separately. Affinity analysis beyond temporal correlation enhances the diagnostics used to determine root cause.

This is particularly significant when a fault or performance issue results in multiple alarm signals being raised in multiple system layers. When affinity is used in these instances, incidents can be grouped together and treated as a single incident with a single ticket being generated. Managing faults and performance across service layers and operational silos drives significant improvements in operational efficiency and service assurance.



Groupings are based on both temporal overlap and affinity scores.

The top cluster of incidents are grouped together because they have both high affinity and temporal overlap.

PROBABLE CAUSE

Several algorithms are used in Probable Cause Analysis. One of the most best uses a ranking score based on the combination of Severity, Eccentricity, and Entropy.

Severity is a measure of degree that a system component or element is detrimentally affected by an incident.

Eccentricity is a measure of “disproportional impact” by an incident, and is determined by comparing the affect of an incident on a given component as compared to its peer components.

Entropy, informally, is a measure of disorder in a system (or component), with a perfectly running system having near 0 entropy and a completely dysfunctional system (or component) having high entropy. Entropy is directly related to the rareness of an event. Faults are rarer than non-faults. Higher severity faults are even rarer still and thus have a higher entropy.

A component that exhibits both high severity and high eccentricity, i.e., is the most disproportionately affected by the fault or performance issue as compared to its peers, has the highest entropy. Formally, Entropy is computed from the joint probability of Severity and Eccentricity and other key indicators. Components with high entropy have very high diagnostic value in determining root cause.

Typical Use Case


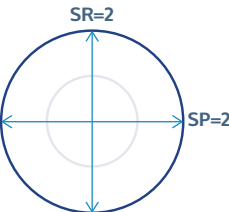

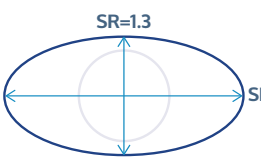
To better understand the roles that Severity, Eccentricity, and Entropy play in determining Probable Cause, let's consider a rather typical use case.

Consider subscribers using a streaming video service. When attempting to watch a piece of content, a high number of users start experiencing “authentication failures” when attempting to connect to their video service using a streaming TV device (e.g., Apple TV, Roku, Chromecast, etc.). The problem could lie anywhere in the system, from the subscriber device to the authentication server in the data center, and all components in between, including the network or the load balancer in front of the authentication service. To correctly diagnose root cause, all components in the topology from the client-device to the authentication service must be analyzed.

Many of those components may be displaying faulty and anomalous behavior, of varying severity, because they are impacted by the underlying failure. To determine probable cause, the fault and metric data from anomalous components must be analyzed at a very fine level, down to the features (dimensions) that characterize the components. For example, client-devices will be analyzed across features such as model and firmware. If all models and all firmware are showing the same severity of authentication failures, then the problem likely lies elsewhere, since it is highly unlikely (but not impossible) for all models and firmware to be faulty at the same time. However, if certain models or firmware are experiencing significantly more severe (higher) fault rates than others, then those clearly become a candidate for probable cause. Even if not the actual cause, the most severely affected components are likely to be a key symptom that points the way to the actual cause. In either case, whether probable cause or key symptom, the severely affected components are of high diagnostic value.

The above example and the relationships between Severity, Eccentricity, Entropy and Diagnostic value can be illustrated geometrically. Let's assume that we are considering whether the Video streaming problem is associated with Roku devices or lies elsewhere.

Examples illustrating Severity, Eccentricity, and Entropy for Roku Devices versus its Peers

SCENARIO		CAUSE	
<p>Normal operation of Roku devices and its peers (e.g., Apple TV, Chrome,...). Note that 1 is normal.</p>	 <p>A circle with a smaller concentric circle inside. A vertical double-headed arrow labeled 'SR=1' spans the diameter of the outer circle. A horizontal double-headed arrow labeled 'SP=1' spans the diameter of the inner circle.</p>	<p>Severity of Roku (SR) = 1 Severity of Peers (SP) = 1 Eccentricity = 1 Entropy ~ 0</p>	
<p>Roku and its Peers are both experiencing an incident of Severity 2.</p>	 <p>A circle with a smaller concentric circle inside. A vertical double-headed arrow labeled 'SR=2' spans the diameter of the outer circle. A horizontal double-headed arrow labeled 'SP=2' spans the diameter of the inner circle.</p>	<p>Severity of Roku (SR) = 2 Severity of Peers (SP) = 2 Eccentricity = 1 Entropy ~ low Diagnostic Value = low Probable cause = unlikely</p>	<p>Roku is not likely a probable cause by itself; instead, most likely it is being affected by the same cause as its peers.</p>
<p>Here Roku is disproportionately affected by an incident relative to its Peers.</p>	 <p>An elongated ellipse with a smaller concentric circle inside. A vertical double-headed arrow labeled 'SR=3.0' spans the major axis of the ellipse. A horizontal double-headed arrow labeled 'SP=1.3' spans the diameter of the inner circle.</p>	<p>Severity of Roku (SR) = 3.0 Severity of Peers (SP) = 1.3 Eccentricity = 2.3 Diagnostic Value ~ Entropy ~ high Probable cause = likely candidate</p>	<p>A high Severity together with a high Eccentricity suggests that Roku is a likely probable cause or a key diagnostic symptom.</p>
<p>Here Roku is less affected by an incident than its Peers.</p>	 <p>A horizontally elongated ellipse with a smaller concentric circle inside. A vertical double-headed arrow labeled 'SR=1.3' spans the minor axis of the ellipse. A horizontal double-headed arrow labeled 'SP=2.6' spans the major axis of the ellipse.</p>	<p>Severity of Roku (SR) = 1.3 Severity of Peers (SP) = 2.6 Eccentricity = 0.5 Diagnostic Value ~ Entropy = medium Probable cause = unlikely</p>	<p>A moderate Severity together with a below normal Eccentricity implies that Roku is not a probable cause, but is a possible key diagnostic symptom.</p>

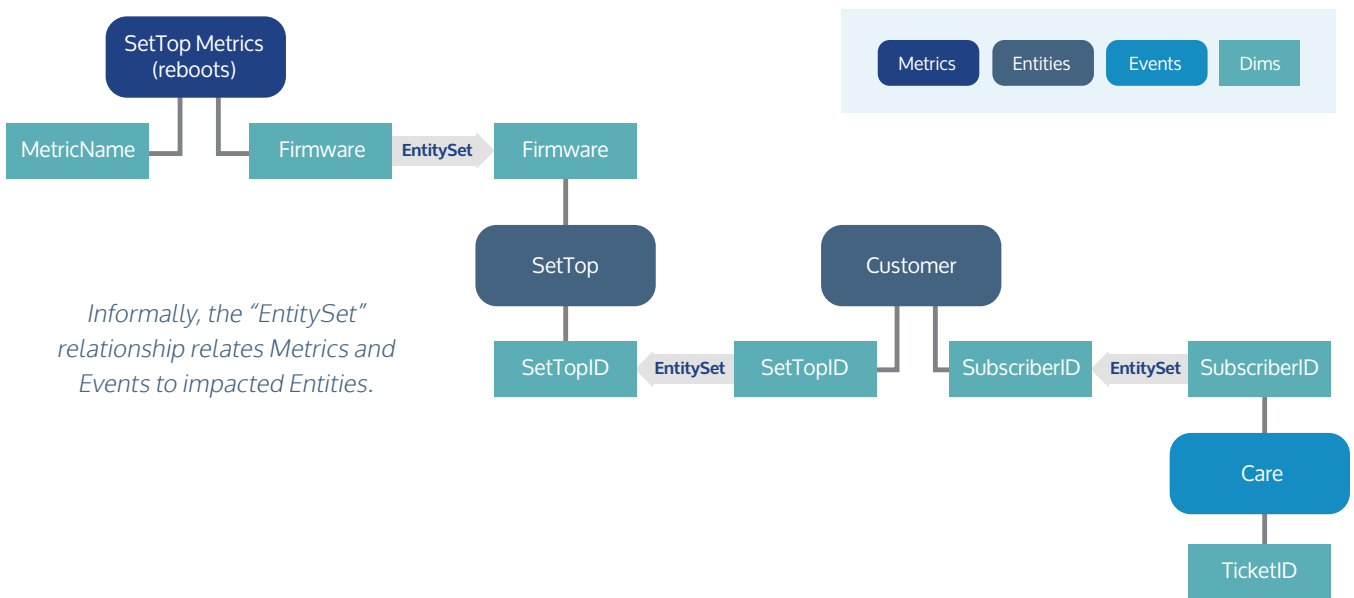
ONTOLOGY REASONING

We define a system to be the composition of all system components and entities being monitored along with the artifacts used in monitoring including the metrics, events, and incidents. An entity may include a device or even a specific customer.

Ontology of a system is information on the logical, topological and physical characteristics across and between devices, infrastructure, customers, and all other system components and entities. Augmented intelligence and machine learning can automatically discover the system ontology. Ontology provides deeper and richer metadata to accelerate automated analysis and diagnosis across the system and subsystems. It can be used in affinity analysis discussed previously to support the grouping of incidents across components and service layers. It can also be used to assess impact and provide deep insight into performance related issues.

In the previous video streaming example, the use of ontology reasoning can determine the extent of the problem, the population impacted and the cost of impact.

Ontology Reasoning Defining Impact



Firmware X is affecting 118,733 Client-Devices. This is 0.427% of all Client-Devices.

This, in turn, is affecting 118,733 (0.213%) Customers.

1,598 affected Customers have recent service-impacting events.

117,135 affected Customers have no recent service-impacting events.

Estimated Cost of current service-impacting events: \$14,382.

Potential Cost of all 118,733 affected Customers: \$1,068,597.

SUMMARY

When considering a next generation AIOps application, data science is a clear differentiator.

How data science is used impacts:

What **processes** and use cases can be optimized.

Their ability to **uncover** root cause and separate cause from symptoms across the service technology stack and across subsystems.

The **insight** provided in order to accelerate not only response but resolution.

The **speed** at which performance issues and faults are detected.

Ask the difficult questions to understand the analytic strengths and weaknesses of AIOps applications prior to your final selection.



[Learn more about VIA AIOps.](#)

ABOUT VIA AIOps

VIA AIOps is a next generation AIOps application that enables intelligent automation across all layers of service delivery to improve the customer experience and optimize operations. VIA AIOps provides total ecosystem observability, and explanatory AI to increase confidence in automation. VIA AIOps delivers noise reduction, correlation, and intelligent automation across operational silos to enhance customer experience and reduce operational cost by enabling more rapid issue detection, mitigation and resolution.

